

Claims

What is claimed is:

1. A method for tracking an object of interest in a video processing system, the method comprising the steps of:

5 generating for a given measurement interval an audio locator output and a video locator output, each indicative of a location of the object of interest;

10 applying a set of rules to determine a manner in which at least one of the audio locator output and the video locator output will be utilized to adjust a setting of the camera based on the given measurement interval; and

15 adjusting the camera setting in accordance with the determined manner of utilization.

20 2. The method of claim 1 wherein the object of interest comprises a moving person.

25 3. The method of claim 1 wherein the camera is a pan-tilt-zoom (PTZ) camera having adjustable pan, tilt and zoom settings.

30 4. The method of claim 1 wherein the set of rules includes determining if the audio locator and video locator outputs are sufficiently close for the given measurement interval, and utilizing only the audio locator output to adjust the camera setting if the audio and video locator outputs are not within a specified range of one another for the given measurement interval.

35 5. The method of claim 4 wherein the set of rules further includes utilizing the video locator output to adjust the camera setting only if the audio and video locator outputs are within a specified range of one another for the given measurement interval.

6. The method of claim 5 wherein the set of rules further includes determining if a confidence indicator associated with the video locator output is above a specified video locator threshold for the given measurement interval, and utilizing the video locator output to adjust the camera setting only if the video locator confidence indicator is above the video locator threshold for the given measurement interval.

7. The method of claim 1 wherein the set of rules includes determining based on the audio locator output if the object of interest corresponds to a new speaker in a multiple-participant system, and if a new speaker is detected, directing the camera to zoom out by a predetermined amount and to turn in a direction of the new speaker.

8. The method of claim 1 wherein the set of rules includes determining based on the audio locator output if the object of interest corresponds to a same speaker in a multiple-participant system, and if a same speaker is detected, utilizing the video locator output to adjust the camera setting so as place the same speaker at a designated position within one or more video frames generated by the camera.

9. The method of claim 8 wherein the set of rules further includes adjusting a zoom setting of the camera until a head of the identified same speaker occupies a designated portion of a given one of the one or more video frames generated by the camera.

10. The method of claim 1 wherein the set of rules specifies that the camera is zoomed out by a predetermined amount after a detected period of continued silence exceeds a first amount of time.

11. The method of claim 10 wherein the set of rules further specifies that the camera is zoomed out by an additional amount if the detected period of continued silence exceeds a second amount of time greater than the first amount of time.

5

12. An apparatus for tracking an object of interest in a video processing system, the apparatus comprising:

a camera; and

10 a processor coupled to the camera and operative (i) to process an audio locator output and a video locator output, each indicative of a location of the object of interest for a given measurement interval; and (ii) to apply a set of rules to determine a manner in which at least one of the audio locator output and the video locator output will be utilized to adjust a setting of the
15 camera based on the given measurement interval, such that the camera setting is adjusted in accordance with the determined manner of utilization.

13. The apparatus of claim 12 wherein the object of interest
20 comprises a moving person.

14. The apparatus of claim 12 wherein the camera is a pan-tilt-zoom (PTZ) camera having adjustable pan, tilt and zoom settings.

25

15. The apparatus of claim 12 wherein the set of rules includes determining if the audio locator and video locator outputs are sufficiently close for the given measurement interval, and utilizing only the audio locator output to adjust the camera
30 setting if the audio and video locator outputs are not within a specified range of one another for the given measurement interval.

16. The apparatus of claim 15 wherein the set of rules further includes utilizing the video locator output to adjust the camera setting only if the audio and video locator outputs are within a specified range of one another for the given measurement interval.

17. The apparatus of claim 16 wherein the set of rules further includes determining if a confidence indicator associated with the video locator output is above a specified video locator threshold for the given measurement interval, and utilizing the video locator output to adjust the camera setting only if the video locator confidence indicator is above the video locator threshold for the given measurement interval.

18. The apparatus of claim 12 wherein the set of rules includes determining based on the audio locator output if the object of interest corresponds to a new speaker in a multiple-participant system, and if a new speaker is detected, directing the camera to zoom out by a predetermined amount and to turn in a direction of the new speaker.

19. The apparatus of claim 12 wherein the set of rules includes determining based on the audio locator output if the object of interest corresponds to a same speaker in a multiple-participant system, and if a same speaker is detected, utilizing the video locator output to adjust the camera setting so as place the same speaker at a designated position within one or more video frames generated by the camera.

20. The apparatus of claim 19 wherein the set of rules further includes adjusting a zoom setting of the camera until a head of the identified same speaker occupies a designated portion of a given one of the one or more video frames.

21. The apparatus of claim 12 wherein the set of rules specifies that the camera is zoomed out by a predetermined amount after a detected period of continued silence exceeds a first amount
5 of time.

22. The apparatus of claim 21 wherein the set of rules further specifies that the camera is zoomed out by an additional amount if the detected period of continued silence exceeds a second
10 amount of time greater than the first amount of time.

23. An article of manufacture comprising a storage medium for storing one or more programs for tracking an object of interest in a video processing system, wherein the one or more programs when
15 executed by a processor implement the steps of:

generating for a given measurement interval an audio locator output and a video locator output, each indicative of a location of the object of interest;

applying a set of rules to determine a manner in which at
20 least one of the audio locator output and the video locator output will be utilized to adjust a setting of the camera based on the given measurement interval; and

adjusting the camera setting in accordance with the determined manner of utilization.